

**Management of multi-scale forest resource data over  
time**

Jussi Rasinmäki

Department of Forest Resource Management  
Faculty of Agriculture and Forestry  
University of Helsinki

Academic dissertation

To be presented with permission of the Faculty of Agriculture and Forestry, University of Helsinki, for public criticism in Walter-hall, EE-building, Agnes Sjöbergin katu 2, Helsinki, on November 9<sup>th</sup>, 2007 at 12 o'clock noon.

*Title of dissertation:* Management of multi-scale forest resource data over time

*Author:* Jussi Rasinmäki

*Dissertationes Forestales 49*

*Thesis supervisors:*

Prof. emeritus Jouko Laasasenaho

Faculty of Agriculture and Forestry, University of Helsinki, Finland

Prof. Timo Tokola

Faculty of Forest Sciences, University of Joensuu, Finland

*Pre-examiners:*

Dr. Keith M. Reynolds

Pacific Northwest Research Station, Forest Service, U.S. Department of Agriculture, USA.

Prof. Ari Jolma

Department of Surveying, Institute of Cartography and Geoinformatics, Helsinki University of Technology, Finland.

*Opponent:*

Prof. Kim Lowell

Cooperative Research Centre for Spatial Information, University of Melbourne and Department of Primary Industries-Victoria, Parkville, Victoria, Australia.

ISSN 1795-7389

ISBN 978-951-651-186-6 (PDF)

(2007)

*Publishers:*

Finnish Society of Forest Science

Finnish Forest Research Institute

Faculty of Agriculture and Forestry of the University of Helsinki

Faculty of Forest Sciences of the University of Joensuu

*Editorial Office:*

Finnish Society of Forest Science

Unioninkatu 40A, 00170 Helsinki, Finland

<http://www.metla.fi/dissertationes>

**Rasinmäki, J.** 2007. Management of multi-scale forest resource data over time. *Dissertationes Forestales* 49. 39 p. Available at <http://www.metla.fi/dissertationes/df49.htm>

During the last decades there has been a global shift in forest management from a focus solely on timber management to ecosystem management that endorses all aspects of forest functions: ecological, economic and social. This has resulted in a shift in paradigm from sustained yield to sustained diversity of values, goods and benefits obtained at the same time, introducing new temporal and spatial scales into forest resource management.

The purpose of the present dissertation was to develop methods that would enable spatial and temporal scales to be introduced into the storage, processing, access and utilization of forest resource data. The methods developed are based on a conceptual view of a forest as a hierarchically nested collection of objects that can have a dynamically changing set of attributes. The temporal aspect of the methods consists of lifetime management for the objects and their attributes and of a temporal succession linking the objects together. Development of the forest resource data processing method concentrated on the extensibility and configurability of the data content and model calculations, allowing for a diverse set of processing operations to be executed using the same framework. The contribution of this dissertation to the utilisation of multi-scale forest resource data lies in the development of a reference data generation method to support forest inventory methods in approaching single-tree resolution.

**Keywords:** spatio-temporal data model, hierarchy, history management

## ACKNOWLEDGEMENTS

I would like to thank my supervisors Prof. Jouko Laasasenaho, especially for igniting my interest in the topic of the thesis, and Prof. Timo Tokola, for the essential "last kicks on the behind" to see the work taken to completion. Numerous discussions with them throughout the years have naturally been instrumental in developing the ideas sed in the work.

Many researchers have been involved in different phases of the work. I would like to thank Sakari Ilomäki for his unfaltering resolution to the tackle the mountain of data thrown at him. Dr. Ilkka Korpela and Timo Melkas showed a commendably determined approach to field data collection and data processing. Big thanks to my long-time office roommate Hanna Huitu for inspirational discussions and the brave efforts to make our room a nicer place to work in. Antti Mäkinen and Jouni Kalliovirta were not only responsible for creating an enjoyable working atmosphere but also for forming two thirds of a research team that it has been a pleasure to be a part of. With Risto Viitala from HAMK University of Applied Sciences I have been able to share ideas, and practical problems, concerning reconstructing the past of a forest area. Prof. Annika Kangas and the preliminary examiners for the thesis, Dr. Keith Reynolds and Prof. Ari Jolma, provided valuable comments for improving the manuscript, while Malcolm Hicks was responsible for making it readable.

This work was funded through several projects. The groundwork was done with the support of the "Forests in GIS" graduate school financed by Metsämiesten säätiö and the "Aika-paikkatieto metsätaloudessa" project financed by Metsähallitus, Metsämannut Oy and the Finnish Funding Agency for Technology and Innovation (Tekes). The work was then continued as part of the "New generation planning system for forest management" project financed by Metsähallitus, Metsämannut Oy, the Forestry Development Centre Tapio, Tornator Oy, UPM-Kymmene Ltd. and Tekes. Support for the final stage was also provided by the "Improving data use efficiency" project funded by the Finnish Academy (project #200775).

Special thanks are due to my parents for never being short of encouragement and nurturing my inquisitive nature as I was growing up. I wish to express my gratitude to my son Aapo for steadfastly dragging me from the world of data management to the realm of steam engines and diggers. The warmest thanks of all go to my wife, Annukka, for always being there with unconditional support.

## LIST OF ORIGINAL ARTICLES

- I. Rasinmäki, J. 2003. Modelling spatio-temporal environmental data. *Environmental Modelling & Software* 18: 877–886.
- II. Rasinmäki, J., Mäkinen, A. & Kalliovirta, J. 2007. An adaptable simulation framework for multi-scale forest resource data. Submitted manuscript.
- III. Rasinmäki, J. 2007. XQuery as a retrieval mechanism for longitudinal multi-scale forest resource data. Submitted manuscript.
- IV. Rasinmäki, J. & Melkas, T. 2005. A method for estimating tree composition and volume using harvester data. *Scandinavian Journal of Forest Research* 20: 85–95.

Rasinmäki was responsible for most of the writing of Papers II and IV. The method and system development reported in Papers II and IV was carried out by all the authors together. For Paper II Rasinmäki was responsible for the implementation of the developed algorithms.

## TABLE OF CONTENTS

1 Introduction.....	8
1.1 Background.....	8
1.2 Multiple scales of forest resource data—operational scale versus grain size.....	10
1.3 Spatio-temporal data management.....	11
1.3.1 <i>Technical aspects of data management</i> .....	12
1.4 Aims.....	12
2 Conceptualisation of a forest.....	13
3 Case material.....	14
3.1 Development of the methods.....	14
3.2 Testing of the methods.....	15
4 Results.....	17
4.1 Effect of spatial scale on forest resource data management.....	17
4.1.1 <i>Storage</i> .....	17
4.1.2 <i>Processing</i> .....	20
4.1.3 <i>Access</i> .....	21
4.1 Effect of time on forest resource data management.....	22
4.2.1 <i>Storage</i> .....	22
4.2.2 <i>Processing</i> .....	23
4.2.3 <i>Access</i> .....	24
4.3 Trees that once were—what do they reveal of the forest that was?.....	27
5 Discussion.....	27
References.....	33



# 1 Introduction

## 1.1 Background

The purpose of this work was to develop methods that would enable spatial and temporal scales to be introduced into the storage, processing and utilisation of forest resource data. Excluding the cartographic scale associated with maps, Lam and Quattrochi (1992) provide the following definitions of spatial scale that can be applied to the temporal dimension as well: (i) scale as the spatial extent of an area to be studied, i.e. geographical scale, domain (Turner et al. 1989), (ii) scale as the smallest distinguishable part of the spatial data set, i.e. resolution, grain (Turner et al. 1989), and (iii) the scale at which a process operates in the environment, i.e. operational scale.

When the management of forest resources is equated with timber management, the domain of forest resource data in the case of Finland has been the forest holding of a single owner, or in case of centralised planning of the management of private forests, a group of forest holdings. The resolution has been the forest stand and the operational scale has varied between the forest holding and the stand depending on the type of forestry operation. The most commonly used operational scale on the temporal dimension has spanned 10 years into the future, while the past has been largely neglected. The temporal resolution of the data itself is one year, while the resolution used with the operational scale can be finer than that, up to one day.

There has been a global shift in forest management during recent decades from a focus solely on timber management to ecosystem management that endorses all aspects of forest functions: ecological, economic and social (Iftekhhar 2005). This has resulted in a shift in paradigm from sustained yield to sustained diversity of the values, goods and benefits obtained (Rauscher and Reynolds 2005). These new objectives introduce new temporal and spatial scales into management.

Two fundamental and interconnected themes in ecology are the development and maintenance of a spatial and temporal pattern and the consequences of that pattern for the dynamics of populations and ecosystems. Central to these questions is the issue of how the scale of observation influences the description of a pattern, as each individual and each species experiences the environment on a unique range of scales, and thus responds to variability in an individualistic manner (Levin 1992). Thus, where the conservation of biodiversity is concerned, one should adjust the scale of observation to the appropriate landscape patches for key species. Biodiversity can also be assessed through species richness, an attribute of broad-scale landscape mosaics. To observe and monitor the dynamics of species richness one would need to expand the temporal scale of observation as well, as the temporal scale of the system increases with the spatial scale (Wiens 1989). Expansion of the temporal scale is not only associated with broad spatial scales, as at the individual species level, however, as the response to habitat changes may be experienced only after a time lag (Hanski and Ovaskainen 2002).

The question of what scales should be included in a combined system of timber and forest ecosystem management can be approached from different viewpoints. One approach is to study the structural characteristics of a forest on different scales simultaneously. The spatial scales combined in these studies vary from sub-basins (thousands of km<sup>2</sup>) to sub-catchment areas (tens of km<sup>2</sup>) (Wimberly and Ohmann 2004), from landscape to stand (Pennanen et al. 2004, Mendoza et al. 2005), from sub-catchment to stand (Reynolds and



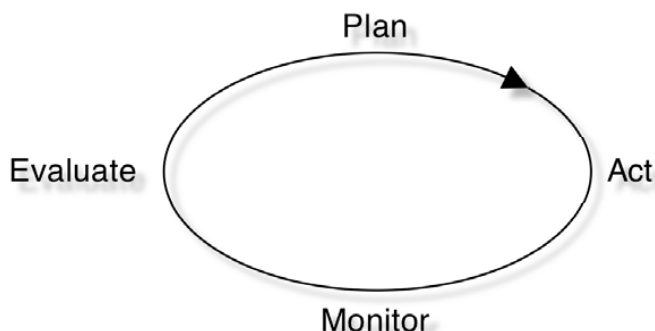
Hessburg 2005), for forest to sample plot (Montes et al. 2005) and as far as within-stand variation (García-Gigorro and Saura 2005, Wolf 2005, Zenner 2005). Another approach is to concentrate on a particular phenomenon and the appropriate scale for measuring it, i.e. scales for measuring forest fragmentation (García-Gigorro and Saura 2005), biodiversity (Yue et al. 2005), the risk of natural disturbance (Barbour et al. 2005), or the socioeconomic services provided by a forest (Horne et al. 2005, Köchli and Brang 2005).

There are only few studies concentrating on the appropriate time scale for measuring changes in structural properties or other forest phenomena. The time scale used in evaluations of changes in the past is usually limited to decades (Wimberly and Ohmann 2004, Montes et al. 2005, Löfman 2006), whereas simulation studies charting possible futures can cover spans of up to 10,000 years (Pennanen et al. 2004).

The question of scale can also be approached from a species point of view. The answer to the question is necessarily species-dependent, as there is no single correct scale on which to describe species-habitat relationships (Wiens 1989), and each species responds to its environment on a range of scales (Levin 1992).

Depending on the characteristics of the species concerned, the relevant scales can vary dramatically. Nams et al. (2006), studying the habitat preferences of grizzly bears, found all the resolutions examined, from 1 km<sup>2</sup> to 6400 km<sup>2</sup>, to be significant, and a similar finding emerged in the case of the capercaillie (*Tetrao urogallus*), but the scale varied from 1 to 1,100 ha (Graf et al. 2004), whereas Kurki et al. (2000) found 100 km<sup>2</sup> to be the spatial extent that best explained the variability of nesting success for both the capercaillie and black grouse (*Tetrao tetrix*). The findings of Moore et al. (2000) emphasize that when evaluating the effect of management actions on a certain species, knowing the correct spatial scales is not enough if the information on populations of the species concerned is incomplete, e.g. only presence/absence data without information on the spatial distribution.

Ecosystem management can be operationalised as a form of adaptive management that consists of a cycle (Fig. 1) of planning, action, monitoring and evaluation (Walters 1986, Reynolds 2005). This allows for the integration of new goals, knowledge and technology into the management at the stage where the results of monitoring the previous round of planning and action is evaluated. This nevertheless places requirements on the ability of the data management system to assimilate modifications and extensions to its data content and processing methods. The scale at which data are acquired and analysed may change be-



**Figure 1.** Phases in the adaptive management cycle (adapted from Maser et al. 1994).

tween successive planning actions and especially in boreal forests, there can be a considerable time lag, quite often measured in decades, between planning and implementation, and also between implementation and evaluation. The past can therefore not be neglected in forest data management, if the goal is to support adaptive forest management. The ability to handle a long time frame and the associated changes in data content gains further in emphasis in the event of a policy failure being detected. In such cases the key variables have not usually been recognised before and new monitoring systems have to be rapidly introduced as a corrective measure (Walters 1986), e.g. forest health monitoring in the 1980s and 1990s (Hall 1995, Ferretti 1997). The compatibility of the new monitoring system with the old one should be ensured in order to maintain an ability to review the changes.

## **1.2 Multiple scales of forest resource data—operational scale versus grain size**

Trees are the basic units of resource information for forest products, and other levels in the information hierarchy can be obtained by aggregation from the tree level. At its simplest and most straightforward the term forest resources in the boreal vegetation zone means trees. Trees are physical objects with well-defined geographical and temporal locations. In terms of planar geometry the spatial location of a tree can be approximated with a point that does not change position over the lifetime of the tree, although viewed in three dimensions, it obviously changes in vertical extent as it grows. If trees were the only objects that forest resource data management deals with, it would be quite straightforward to define their spatial and temporal extents by stating their point location and lifetime. The time dimension would even shrink to an instant and could thus be ignored if the resource manager were only interested in the current situation in the forest, a common practice until recently apart from plans to retain a given cutting volume over a longer period of time.

The transition from the basic operational scale of a single tree to the data collection grain size of a single tree is anything but trivial, however. The compiling of an inventory of all the trees in a forest and data management at that level have been impaired because of both impracticalities associated with obtaining data on individual trees and the demands that such a task would impose on data processing. For cost reasons, all the conventional methods of producing forest resource data are based on samples drawn from a complete enumeration of the trees in the target area. These samples are then aggregated to predict the properties of the forests delineated in the area. The sampling scheme is by necessity tied to the spatial scale of the delineation in order to achieve the required information in the most cost-efficient way. The delineations are based either on the properties of the set of trees, i.e. the forest stand, or based on the location of the trees, e.g. forest resource data derived from satellite images and managed in units that correspond to the spatial extent of an image pixel on the ground. In both cases data management operates on only one spatial scale, as even in stand-based data management the stands are quite uniform in size. The fact that stand delineation is based both on the requirements of forestry operations and on the characteristics of the trees and soil leads to a roughly uniform stand size for any given area. The ability to change the temporal or spatial scale of observation has thus not traditionally played much of a role in the production and management of forest resource data, except for the estate-stand dichotomy, where the estate, or region, level has been used to define goals for forest management and the stand or inventory unit level has provided the data. A change to a finer scale of observation would be impractical, as the data sample would not be optimal for the new scale. Were we able to enumerate all the trees in the area under study and state their

locations, changing the spatial scale of observation would not pose a problem if the tree level were indeed the finest scale of interest. The change of scale would just mean a new aggregation of the underlying tree data. Such aggregations could be based on widely different criteria, however, or entail a subjective component, depending on their purpose. Objects at higher levels of aggregation would therefore be likely to lack the clear definition of geographical and temporal location that trees possess.

New forest inventory methods generally strive towards the tree-level scale in order to describe the structure of the forest more accurately and thus tally the timber resources more accurately. As a result, there are inventory methods that aim to approach the tree level by using sub-stand partitions (Hyvönen et al. 2005) or even aim at the complete enumeration of trees. Even the established field inventories which produce aggregated data on partitions of the area concerned usually produce initial data on finer scale within the partitions on which the aggregation is based: sample plots and perhaps trees within the plots. Field data collection techniques have been enhanced recently by facilitating the measurement of single trees, including their exact location (Kalliovirta et al. 2005), and remote sensing methodology and the sensors themselves have been developed in the same direction (e.g. Pekkarinen 2002, Maltamo et al. 2004, Korpela and Tokola 2006). The varying height structure of a natural forest stand as opposed to a plantation forest makes the location of all trees a demanding task, however, when the data signal is captured above the forest. In a heterogeneous stand more than half of the trees can go undetected (Maltamo et al. 2004, Korpela and Tokola 2006).

An alternative approach would be to utilise tree data from other forestry operations, e.g. clear cutting, leading to a dramatic decrease in data acquisition costs. The information content may not be optimal from the inventory point of view, but it may be sufficient, especially as a reference data source to be combined with inventory methods (Stendahl and Dahlin 2002).

### 1.3 Spatio-temporal data management

El-Geresy and Jones (2000) give a classification of models used to link time and space. The main entities of features are their state, their relations to space and time, and their interrelations within space and time. The problem therefore has three dimensions: spatial, temporal and feature. An elementary solution would be to select one of these as a basis for modelling, resulting in location, feature, or time-based models. Concentrating on a single dimension does not facilitate the analysis of the relations between the three, however. There are a number of approaches available for integrating all three dimensions into a single model (Langran 1991, Worboys 1994, Peuquet and Duan 1995, Yuan 1999, Wachowicz 1999, Griffiths et al. 2004). In event-based models, a change between two successive states or locations of objects is represented as an event which links location, feature and time. An alternative approach for viewing changes in state would be to promote the process to the role of a feature; i.e. instead of tracking the states of the system, one represents and stores the processes (Reitsma and Albrecht 2005).

Since the main focus has been on the development of data models, spatio-temporal data access methods have received less attention. One exception is the Tripod model (Griffiths et al. 2004), which extends the ROSE algebra (Güting and Schneider 1995) to the temporal domain and provides a base for implementing a query processing architecture and object manipulation capabilities. Mountrakis et al. (2002) present a difference-based change-

oriented model complemented with the operators needed to aggregate the change over time and to propagate the change across the different resolutions used in the system. Camossi et al. (2006) present an object-oriented framework that explicitly supports conversions between different spatial and temporal granularities.

### *1.3.1 Technical aspects of data management*

Three technical solutions of data storage can be distinguished, depending on how they treat the actual data content and the metadata, i.e. the description of the content of the data. In the file-based approach the raw data is the only stored component, the metadata content being external to the data file. A structured document, while usually also stored as a data file, combines the description of the content with the actual content. The idea of a database is to add a layer of separation between the description of the data and the data as such. The data content can be accessed through the metadata layer without knowledge of the actual implementation of data storage. For database storage there is also a continuum from the decomposition of data into individual attributes, i.e. relational databases, to the storage of objects that are described not only through their attributes properties but also through their behaviour, i.e. object databases.

Once stored, it should be possible to access the data. Data access methods fall into two categories: programmatic access and query language access. The former is the only option for raw data stored in files, as in that case the formal language description of the program will contain the metadata description of the data content. Query languages, e.g. SQL (Chamberlin and Boyce 1974) for relational databases and XQuery (Boag et al. 2005) for structured documents, require existing metadata which they can rely on for data access implementation and which will provide an additional layer of abstraction between how the data is stored and how it is accessed.

## **1.4 Aims**

When viewing adaptive forest resource management as a means of reacting to any undesirable changes in the managed natural environment, the abilities to track how the forest has changed over an extensive period of time and to assess what may have been the driving forces behind the changes arise as crucial aspects. It should also be noted that a forest in this sense can mean a small collection of trees, a landscape of thousands of hectares or anything between, so that the objects recognised in it may not be unambiguous in their geometry and the data collected from it is likely to change considerably in content during the monitoring period.

The aim of the present work was to develop methods that could be used to manage multi-scale forest resource data on both the spatial and temporal dimensions. Data management is defined here as including both the storage and processing of data and access to the stored data. The effects of introducing the spatial and temporal scales into the management of forest resource data were studied with respect to five aspects:

- i) Development of a conceptual data model for the storage of forest resource data on temporal and spatial scales (Paper I)
- ii) Processing forest resource data that spans different scales of spatial observation, concentrating on the extensibility and configurability of the data content and processing tasks (Paper II)

iii) Exploring mechanisms for accessing a multi-scale forest resource data store (Paper III).

iv) Development of a method for generating reference data to support forest inventory methods that approach single tree resolution (Paper IV).

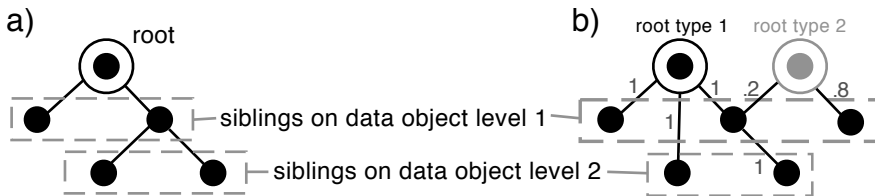
v) Assessment of the suitability of the resulting methods for handling a more general definition of spatial hierarchy than that used in developing the method, including incomplete spatial nesting and uncertain relationships.

## 2 Conceptualisation of a forest

For the purposes of this study, a forest was defined as a collection of objects, trees being the basic objects which are then used to form higher-level objects by aggregation. At the higher level there are also objects that are not derived from the properties of trees but from the properties of the soil and terrain. These are crucial to the ecosystem-based view of forest management and should thus be incorporated into data management. As they are not directly derived from trees, however, their inclusion breaks the assumption of complete spatial nesting of data hierarchy levels within each other. There are also objects at finer levels than that of trees that are relevant to ecosystem management. This microhabitat-level information can be partly derived from the tree-level data, but is otherwise currently unobtainable for any substantial area.

The multilevel nature of the data was expressed as a hierarchy of objects, which was expected from the method development perspective to form a rooted tree graph, i.e. a simple, undirected, connected, acyclic graph with a special root node. Furthermore, all the nodes in the graph were associated with a data object level, e.g. stand-stratum-tree, and the children of any node in the tree were expected to be at the same data level (Fig. 2a). This assumption was relaxed in the summary in order to study the effect of an incomplete hierarchy on the methods developed, i.e. a node in the hierarchy could have children at several data levels (Fig. 2b). This would be beneficial for cases in which the aggregation level depths vary spatially, e.g. a stand-substand-stratum-tree view of a forest in which the stand would have an additional sub-stand child level if the tree composition were to vary considerably within the stand, but at grain sizes that are deemed too small from the operational point of view to warrant stand-level status.

Furthermore, the parentage link was allowed to be uncertain in the summary, i.e. the edges of the graph were labelled with a membership grade between 0 and 1, thus expressing the degree of certainty of the existence of the parental link (Fig. 2b). This kind of uncer-



**Figure 2.** Object hierarchy for expressing spatial scales: a) rooted tree with children at the same descendant level, b) a forest of trees with labelled edges having children at different descendant levels.

tainty may stem from either the spatial or the temporal relationship between data levels. The spatial case arises when two object delineations of different origin are integrated into the same data graph, e.g. an operational stand delineation and an administrative recreational area delineation, and the temporal case when the basis of object delineation changes between two successive instants in time, leading to the spatially inconsistent representation of objects at the same location. For example, an area consisting of a single forest stand may later consist of two stands without any changes in the underlying properties of the forest, simply because the person responsible for the delineation has changed (see Thierry and Lowell 2001).

The whole data structure in the summary was also allowed to have more than one type of top-level object, thus creating a forest of tree graphs (Fig. 2b) reflecting the fact that the conceptualisation of a forest just as a collection of trees is a gross simplification that is not necessarily valid from the ecosystem management perspective. It would be desirable to integrate a host of other data levels not derived from tree data into the data management solution. By allowing the graph to have several roots, more general interdependences between objects at different data levels can be modelled.

### 3 Case material

#### 3.1 Development of the methods

Three data-sets were used to develop the methods presented in papers I, III and IV.

The starting point for the development of a conceptual data model that enables spatially explicit storage of forest resource information over time and at various scales (paper I) was a data-set covering 4,900 hectares of boreal forest in Central Finland for a period of 75 years (1926–2001). The data had been collected using stand-wise forest inventory techniques as part of the monitoring of the use made of state forests. The spatial form of the data was a subdivision of the whole forested area into adjacent polygons, each representing a homogeneous area of the forest with regard to species composition, size structure of the trees and factors such as soil type. The data-set contained observations from six points in time: 1926, 1936, 1949, 1962, 1988 and 2001 and representing a single level, that of the stand, for the data-sets up to 1962 . and a two-level hierarchy with stand and tree species-sized strata from 1988 onwards.

Data from nine inventories carried out on a 1,000-hectare forest estate in Southern Finland in 1871, 1907, 1925, 1935, 1950, 1962, 1976, 1984 and 1994 were used to illustrate the data retrieval method (paper III). The stand linkages between the points in time had been determined manually by the composer of the data-set, employing attribute information and spatial overlay techniques.

A complete census of trees on two final felling stands in Central Finland was used to develop the sub-stand tree composition and volume reference prediction method presented in paper IV. A total of 1,474 trees occupying a total area of 3.2 ha were surveyed using a tachymeter, and the resulting data were combined with the individual tree measurement data recorded by the harvester. To enable combination of the data, the order in which the trees were felled, locations of the harvester and the time when each tree was felled were also recorded.

### 3.2 Testing of the methods

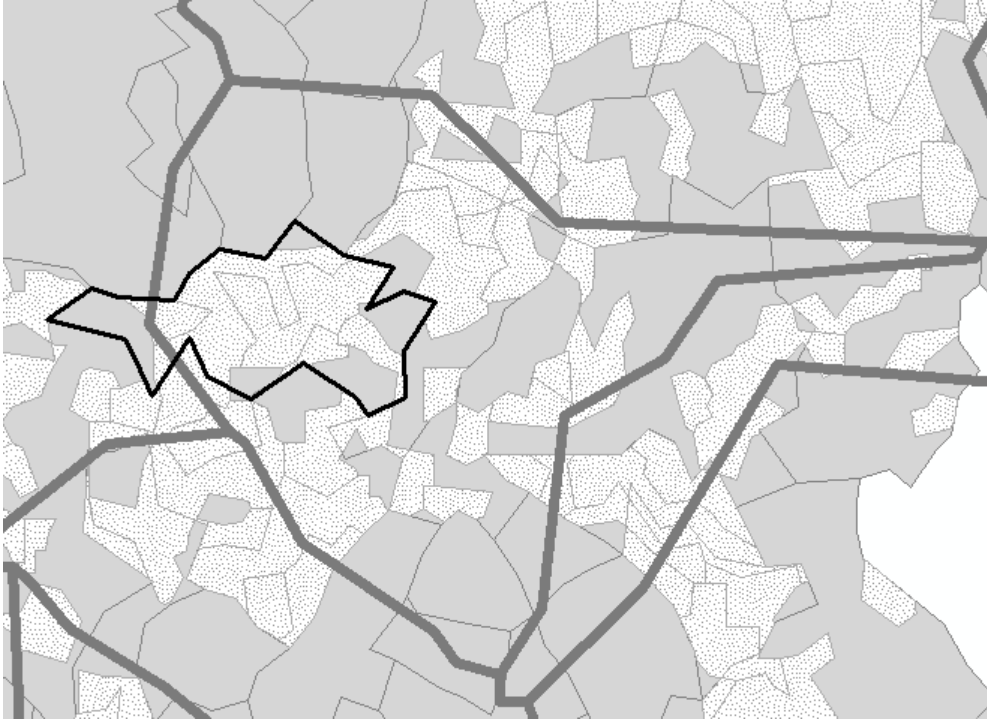
To illustrate the methods developed here and to study the effects of the relaxed data structure assumptions on them, an artificial forest data-set was synthesised that aimed to mimic adaptive forest management data that consist of timber data at a far more detailed level than is customary nowadays and to integrate non-timber data sources as well. In this case soil and catchment area data were used as ‘proxies’ for ecological factors. The timber resource data were incompletely nested at three levels, the top, stand level representing operational units, an optional middle level representing homogeneous areas with regard to tree size and species within the stands, and the lowest level, that of individual trees.

Time is handled in two ways in these methods: as explicit time periods and instants, and as temporal relationships. Since the relaxed assumptions did not affect the explicit representation of time, and since the temporal relationships are conceptually identical to the spatial relationships, there was no specific time component in the test data-set that was generated. The events that caused the changes in the data could also be left out of the test set, as they were either conceptually identical to the spatio-temporal objects or simplified versions of them.

The soil classification delineations produced by the Geological Survey of Finland for a 400 km<sup>2</sup> area in central Finland was selected as a starting point, and a parallel set of top-level objects, catchment areas, was generated from a 25\*25 m elevation model supplied by the National Land Survey of Finland. The catchment areas were generated using the Hydrology Modeling tool for ArcGIS (Hydrology modelling... 2001) with a minimum catchment area of 62.5 ha.

The first child object level for the soil and catchment units in the synthetic data hierarchy consisted of forest stand polygons. The initial stands were generated with the SIMMAP program (Saura and Martínez-Millán 2000) as a clustered landscape of 2,000\*2,000 pixels that were ordered into 3 classes using a p-value of 0.5 and the minimum mapped unit size of 99 pixels. The classified raster was then georeferenced to cover the same 20\*20 km area as the soil class map and converted to polygons. The initial stands were intersected with the soil class objects, excluding the water body polygons. Any resulting intersection polygons that had disjoint parts were disintegrated to single-part polygons, after which polygons smaller than 5,000 m<sup>2</sup> were merged with the neighbouring with which they shared the longest common border. As the stand polygons were created independently of the catchment areas, the spatial nesting of catchments and forest stand polygons was not complete in the sense that a single stand polygon could intersect more than one catchment polygon. The proportion of the whole stand area that overlapped any particular catchment unit was used as the edge label in the data graph expressing the spatial relationship between objects.

SIMMAP was used again to create a relatively unclustered landscape of 2,000\*2,000 pixels containing three classes (p-value 0.3, minimum mapped unit size 10 pixels). The same raster to vector conversion as for the stand layer was performed to obtain a layer of polygons representing the sub-stand level variation in tree composition. A subset of the soil classes was selected to represent those areas in which the tree composition varies on finer spatial scales. The polygons representing those classes were then used as a mask to select only those substand-level variation polygons that intersected with the mask polygons. After studying the intersections between the mask and sub-stand polygons, any resulting polygons smaller than 3,000 m<sup>2</sup> were removed, thus leaving a complete spatial nesting relationship between the stand and sub-stand polygons.



**Figure 3.** A view of the synthetic forest with catchment areas outlined with thick grey lines, stand polygons filled with solid grey and sub-stands with a dotted pattern. One stand crossing a catchment boundary is highlighted with a black border.

The tree data level was generated using a random point pattern. For each stand in the data-set an initial tree set was first generated which covered the minimum bounding box around the polygon with a random tree density between 600 and 2,000 trees/ha. Of the initial tree set the trees that were located inside the stand polygon were selected, and the relationship between the catchment units and possible sub-stands determined. Apart from coordinates, each tree was also allocated a random tree species, height and diameter at breast height.

The total synthetic forest data-set (Fig. 3) consisted of 434 catchments, 3,351 soil units, 14,726 stands, 4,449 sub-stands, and 47,693,031 trees covering the 372 km<sup>2</sup> of land within the 20\*20 km area. The data-set was processed with a program written in the Python language and the data were stored both as serialised Python objects and as XML documents using an Oracle Berkeley DB embeddable database engine (Oracle Berkeley... 2007).



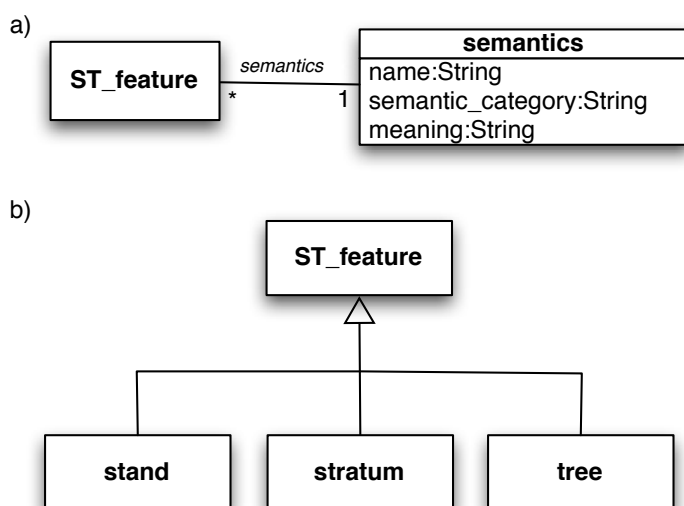
## 4 Results

### 4.1 Effect of spatial scale on forest resource data management

#### 4.1.1 Storage

The conceptual data model presented here (paper I) is based on a single logical entity, a spatio-temporal feature, that has scale-related relations in space. The features form a nested spatial hierarchy, i.e. a top-level feature has child features that can again have children of their own. These spatial parent-child relations are expressed as relational links between the features, corresponding to the edges in a graph representation. The key mechanism for attribute representation is the association of an interpretation with each value. This revokes the need for a fixed data model for feature attributes, but requires the handling of attribute semantics as data. Since the attribute interpretations are no longer fixed in the schema of the data model, a separate semantic class is used to track the attributes within all the data levels. The same mechanism can be used to track the interpretation of each spatio-temporal feature should it not be indicated at the class level, i.e. if all the features are stored in one class (Fig. 4a) instead of separate subclasses of spatio-temporal features for each feature type (Fig. 4b).

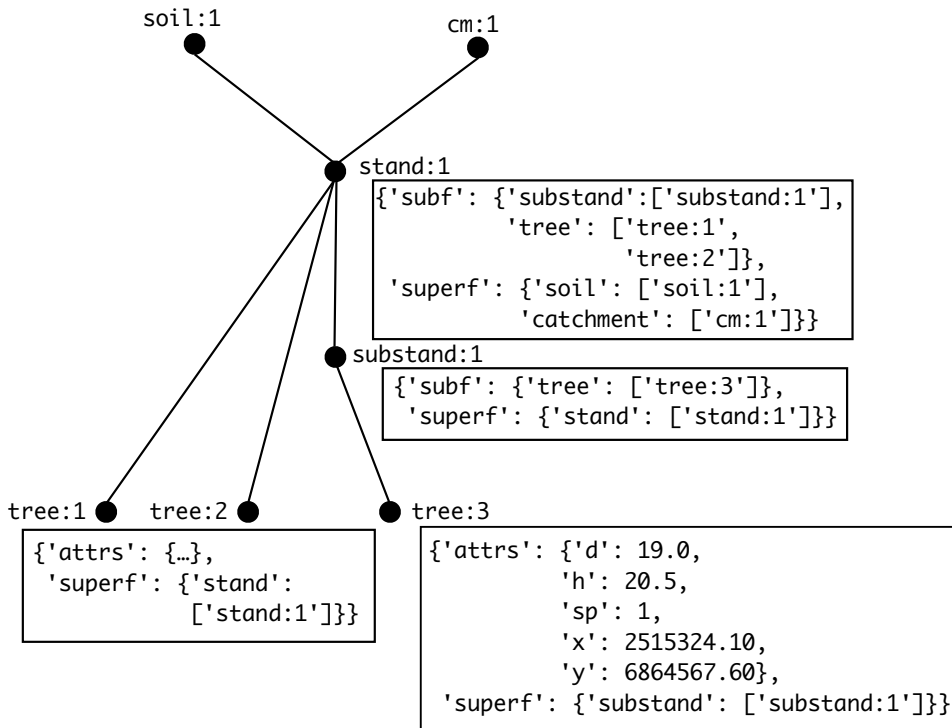
Although it would be possible to implement the conceptual data model in a relational database system, this would be very cumbersome, on account of the hierarchical nature of the model, as relational systems are based solely on set operations (Peuquet 2002), which do not as such support the nesting of entities. Relatively soon after the introduction of the basic relational model (Codd 1970), however, abstraction mechanisms were added that allowed for hierarchical content (Codd 1979), and thus an object-relational database imple-



**Figure 4.** a) A single class for all types of features with an associated semantic class giving the class instance interpretation, b) sub-classes of feature types derived from a common super-class.

mentation of the conceptual model is presented in paper I in which these extensions are employed as well as object-oriented features. The object-relational representation of the database allows the database tables to be viewed as collections of object instances that can have methods in addition to attributes. Two mechanisms for handling spatial nesting are explored in paper III that utilise the hierarchical nature of the Extensible Markup Language XML (Bray et al. 2004). The first mechanism expresses the spatial hierarchy as a direct embedding of objects within each other in the XML document, but this allows only for navigation in a top-down direction for any fragment of the data tree. The second option is identical to the original solution presented in paper I, in which the edge information detailing the connection between two nodes is explicitly stored for each node in the form of either object identifiers or object references.

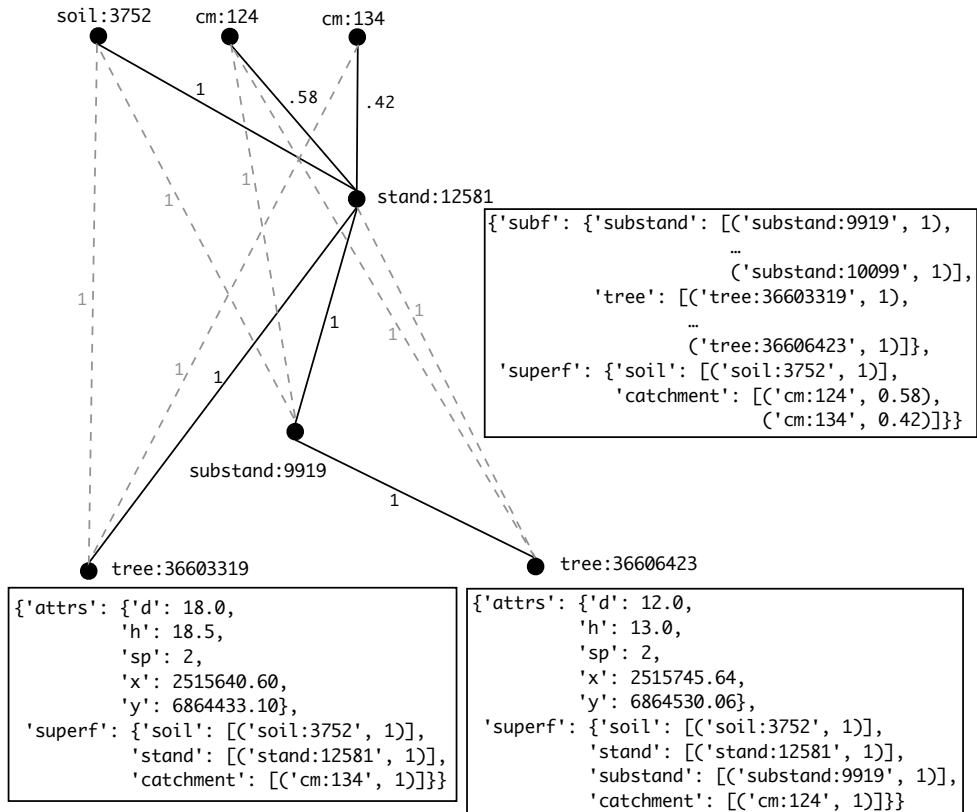
The extension of the data structure from a rooted tree to a forest of trees with non-uniform children structure requires no changes to the data model if the edge information is stored explicitly. The implementation used for the synthetic data-set (Fig. 5) stored the node data as a dictionary or associative array in which the parent and child edges were stored as dictionaries having the parent or child level as their key and a sequence of level node ids as data.



**Figure 5.** Multi-scale data as a forest of trees with non-uniform descendant structure, and the corresponding Python object instances for some of the tree nodes. Two data structures are used: dictionaries are used to give key-value pairs, and lists contain lists of values. The syntax for a dictionary is {key1:value1, key2:value2, ...}. The value in the key-value pair can be of any data structure, including dictionaries and lists. The syntax for a list is [value1, value2, ...].

In the synthetic data-set the edges between the catchment and stand nodes are subject to labelling, which expresses the magnitude of the association between the nodes. The edge labelling can be implemented by storing the label together with the edge. At the same time, however, the number of nodes must be increased to express the relationship between all nodes at all levels, so as not to lose known information, e.g. even though the relationship between a catchment and a stand may be ambiguous, the relationship between the trees in the stand and the catchment is not (Fig. 6).

The whole data-set took 18.3 GB of disk space when stored as serialised Python objects compressed using the zlib algorithm (zlib... 1996). The data were stored in edge-labelled connected graph form, i.e. containing the relationship of any given data node to all data levels. If we assume that there are 77 billion trees taller than 1.3 m in Finland, on 20 million hectares of forest land (Pitkänen 2006), a data-set containing the coordinates, species, height and diameter at breast height for every tree together with higher level aggregated objects and a complete linking of objects throughout the data levels would consume ap-



**Figure 6.** A simple, connected graph with edge labelling, and its Python implementation. Edges express the complete mapping of nodes at different levels, while the edge labels are used to express the magnitudes of the associations between nodes. A further data structure, tuple, is used to define the edges together with their labels. The syntax for a tuple of values is (value1, value2, ...). The parent-child edges are drawn with a solid line and the ancestor-descendant edges with dashed lines.

proximately 30 TB of disk storage when serialised as Python objects in a compressed format.

The first 100 soil units out of the total of 3,351 units, having as their sub-features 368 stands, 6 sub-stands and 1,208,243 trees, took 281 MB of disc space to store in an XML database, and 25 MB as uncompressed serialised Python objects with the same information content as in the XML formulation.

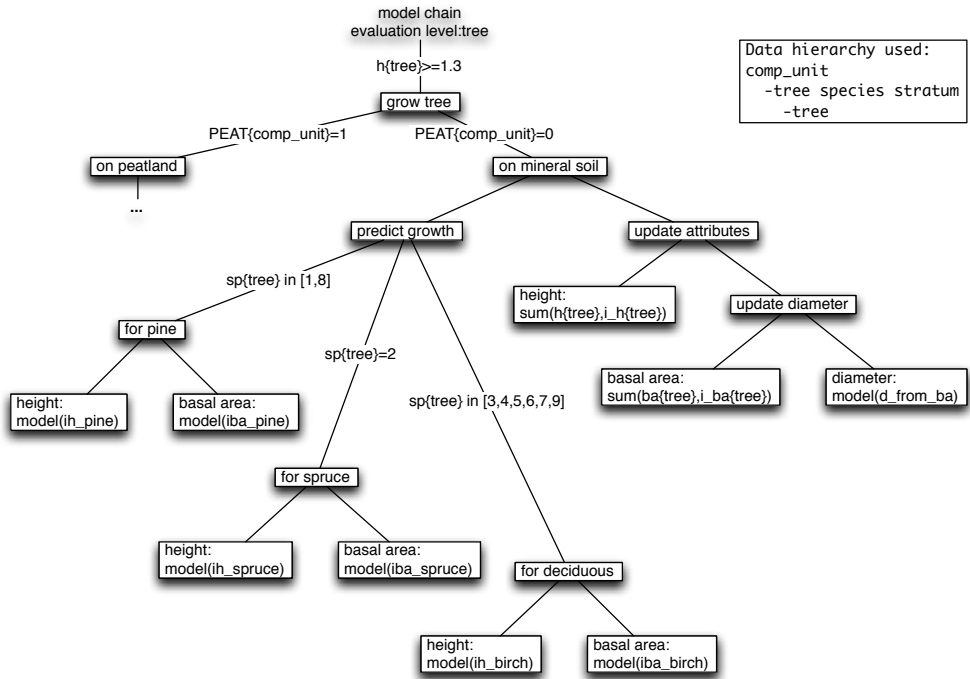
#### *4.1.2 Processing*

The processing of forest resource data is defined here as consisting of two tasks: prediction of non-measured properties for objects and creation of new objects based on known ones and their properties. In planning tasks the models are applied to data in order to predict changes in the properties of features with time.

A forest resource data processing tool is presented in paper II that aims to achieve independence in the processing task, i.e. it should function as a generic forest resource data processing framework. The main design principles were data content configurability, processing model extensibility, and an adaptable control mechanism between the data and the models.

Data content configurability was implemented using a hierarchical data structure that entailed a recursive structure of objects having attributes and sub-objects, where the sub-objects reproduced the basic structure of objects. The attributes at all levels were indicated as value-value interpretation pairs, for which the semantics is expressed as a descriptive domain-specific ontology (Agarwal 2005). In the ontology the different object levels in data hierarchy are given their meaning, i.e. the ontology may state that the topmost object level is the stand, which contains strata objects, which in turn contain trees. The permitted attributes for each object level in the ontology are then stated.

Model extensibility was approached by separating the implementations of the models from the core processing system to form model libraries which shared a common programmatic interface. The individual models were tied to the framework through the same ontology that was used to specify the data. The data processing tasks were formulated as model chains, i.e. sequences of models, each of which was used to predict a certain attribute of the forest. A data structure for implementing the model execution logic was introduced for the model chains, i.e. the processing task logic was not treated as program code but as data. The data structure used to decouple the processing logic from the program code included the composition, order and execution logic of individual tasks in the form of hierarchical task structures (Fig. 7). The tasks describe how individual models can be applied to data to obtain the desired result. The execution of a task may be conditional, and each task may have subtasks that have the same task structure. The leaves of the resulting recursive task tree are formed by the models that ultimately execute the task. A model chain consists of a group of top-level tasks that are assigned to a certain level in the data tree, i.e. the tasks in any model chain only process data at a single level in the data hierarchy, although the input values for the models in the chain may be taken from any level in the data tree, as their parentage and descendants are always known. Thus a data processing operation consists of a data set, the data processing models and a group of model chains for different data levels that have an execution order. By employing a simple, connected, edge-labelled graph as the internal data structure, the data processing tool is able to assimilate and process data that have a non-uniform hierarchy of scale levels and parallel object types on any level.



**Figure 7.** Illustration of the model chain concept in tree graph form: decomposition of tasks into sub-tasks, having models as leaves. The tree is traversed from top to bottom and from left to right, and the traversal conditions are given as edge labels. Note the use of a higher-level attribute (PEAT{comp\_unit}) in the tree-level model chain.

#### 4.1.3 Access

A matching data access mechanism for the conceptual model presented in paper I is crucial, as data access using mechanisms geared towards the fixed-semantics tuple-oriented view of data, i.e. the relational model, is not feasible. The hierarchical data with varying semantics can be accessed using either programmatic implementation together with the storage mechanism or a query language which has support for data of a hierarchical nature.

The implications of using XQuery, the query language for the hierarchical data structure of XML, for data access purposes was analysed in paper III. Various aspects of retrieval from a data store conforming to the hierarchical entity-attribute-value data model were studied by dividing retrieval into three elementary tasks: pivoting, sequencing and aggregation (Johnson and Chatziantoniou 1999). Pivoting serves a convenience function in transforming the attribute-value pairs to a more compact and commonly used columnwise format in which the attribute definitions are detached from the values to serve as column headers in a data table. This transformation also allows tuple-oriented data retrieval mechanisms to be used. The sequencing function is crucial for traversing the hierarchical structures in data, as it can be used to traverse both the temporal lineage of an object and the different spatial scales of the data. Aggregation is used to change the scale of examination of attributes to a coarser level, i.e. when summarising the data content.

These top-level data retrieval tasks were successfully formulated using the For-Let-Where-Order by-Return (FLWOR) syntax and the expression extension capability of XQuery. It was also demonstrated that the nested expression structure of XQuery allows for combining top-level data access tasks to formulate more complex access patterns.

Should the data storage implementation allow it, programmatic access methods mirroring the logic of the XQuery expressions presented here could be formulated. The main difference between the two approaches is that XQuery provides an additional abstraction layer between data storage and access, while programmatic access provides data manipulation capabilities in addition to data retrieval.

## **4.2 Effect of time on forest resource data management**

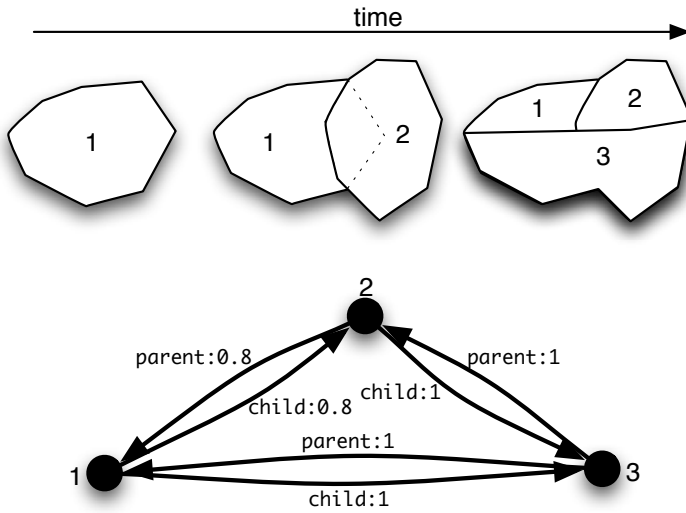
### *4.2.1 Storage*

The effect of time was expressed in internal changes in the objects and the temporal relations between them. The test data in paper I were analysed for changes in both the measurement information content and the definition of a forest stand. The ability of the model to handle these two types of change was the primary objective when designing the data model.

To track internal changes in objects, the spatio-temporal object has a lifetime during which several attributes, in turn having lifetimes of their own, can be associated with it. The value–value interpretation attribute representation coupled with the semantic class allows for temporal change in the set of properties recorded for an object. Along with the names of attributes and their interpretations, the semantic class also stores the lifetime of each attribute within the system. Changes in the spatial representation of an object during its lifetime were stored using the same method, i.e. recognising that an object may undergo several geometrical representations, each of which has a lifetime.

The temporal relations between objects were handled using an identical method to that applied to the spatial relations: regarding the relations as edges in a time graph of object nodes forming links to other objects that are either predecessors or successors of the object in question. The temporal succession of features may be subject to uncertainty should the spatial delineation of objects contain inaccuracies or if the delineation principles are changed. Uncertainty in temporal relations between objects may be expressed using edge labels in an identical fashion to the method used for expressing ambiguity in spatial relations between objects. This would result in a directed, edge-labelled multigraph (Fig. 8), in which the multiple directed edges between nodes will provide for an interpretation of the passage of time in the absence of an identifiable root node for the graph, i.e. for spatial graphs the top-level object, or root node, will anchor the interpretation of the order of spatial levels, but for temporal graphs each edge has to carry information as to whether time is being traversed in the ‘past’ or ‘future’ direction for that particular edge.

To facilitate the analysis of change over time, the data model contains a class for events that drive the changes in spatio-temporal features. In the original model the events occur at points in time, they have locations and have interpretations that are expressed in relation to the semantic class. Events and features are connected through bi-directional relationships. Each feature stores references to the events that have affected it, and each event has a list of references that it has affected. A more detailed relationship between events and features can be computed from the stored locations, i.e. the exact spatial intersection of an event and a



**Figure 8.** Conceptual representation of the temporal relationships between four forest stands, and the associated graph representation. The directed edges represent the direction of temporal lineage: parent or child. The certainty of the existence of an edge in the graph is given as a number between 0 (non-existent) and 1 (existing link) in the edge label.

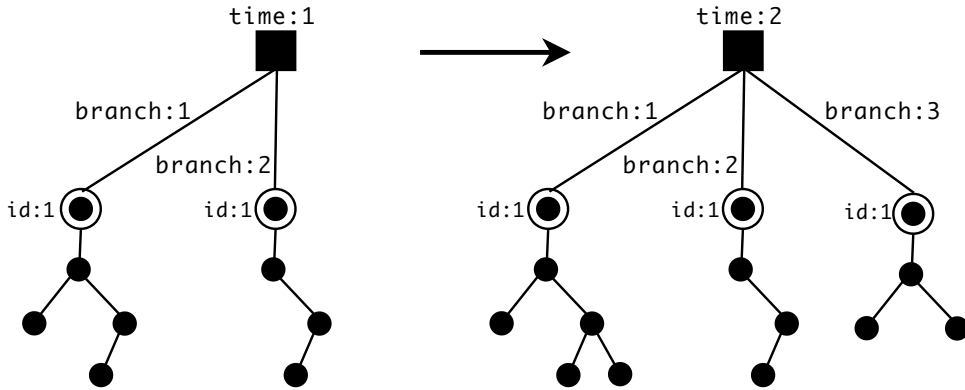
features can be derived from the geometries between the event and the features it has been linked to.

Should the evolution of an event over time be of interest, e.g. the development of a forest fire or insect outbreak, events can also be modelled as a sub-class of spatio-temporal features, thus allowing them to occur over a period of time within which they can have multiple geometries and attribute values.

#### 4.2.2 Processing

Time and the processing of forest resource data are intertwined when models are applied to data in order to predict changes in the properties of features with time. Should the forward, or in rare cases backward, prediction be deterministic, i.e. only a single development path is predicted, the composition of the conceptual model as presented so far will be capable of handling the results, but it is usually desirable in forest management planning to generate several alternative future scenarios for the development of forests based on different management measures and their timing. The measures to be carried out can then be chosen from this set based on the goals and constraints set for forest management. Similarly, the effect of cumulative prediction errors in forest growth simulations can be studied by predicting a set of outcomes using Monte Carlo simulation (see Kangas 1997).

To enable an approach of this kind, an alternative to the original method for time management was presented in paper III. Instead of being centred around objects, the method is time-centric. The results of a simulation are stored for each point in time used in the prediction, as also is the status of the whole data hierarchy. Alternative development paths are stored by branching the data tree at the top level (Fig. 9).



**Figure 9.** The implementation of branching time, a time point-based approach to alternative scenarios.

#### 4.2.3 Access

Temporal data retrieval methods rely on operations that can be used to deduce the relative temporal positions of objects and their temporal succession. In the object-oriented implementation of the conceptual model in paper I, the relative temporal operation was a method associated with an object class for a given period in time. The method supported the comparison of a point in time with a period as well as of two time periods, giving the result as one of 13 temporal relationships (Allen 1983). A similar implementation but limited only to three of the relationships between a point in time and a period, before, within and after, was also presented in paper III, employing the extension mechanism of XQuery.

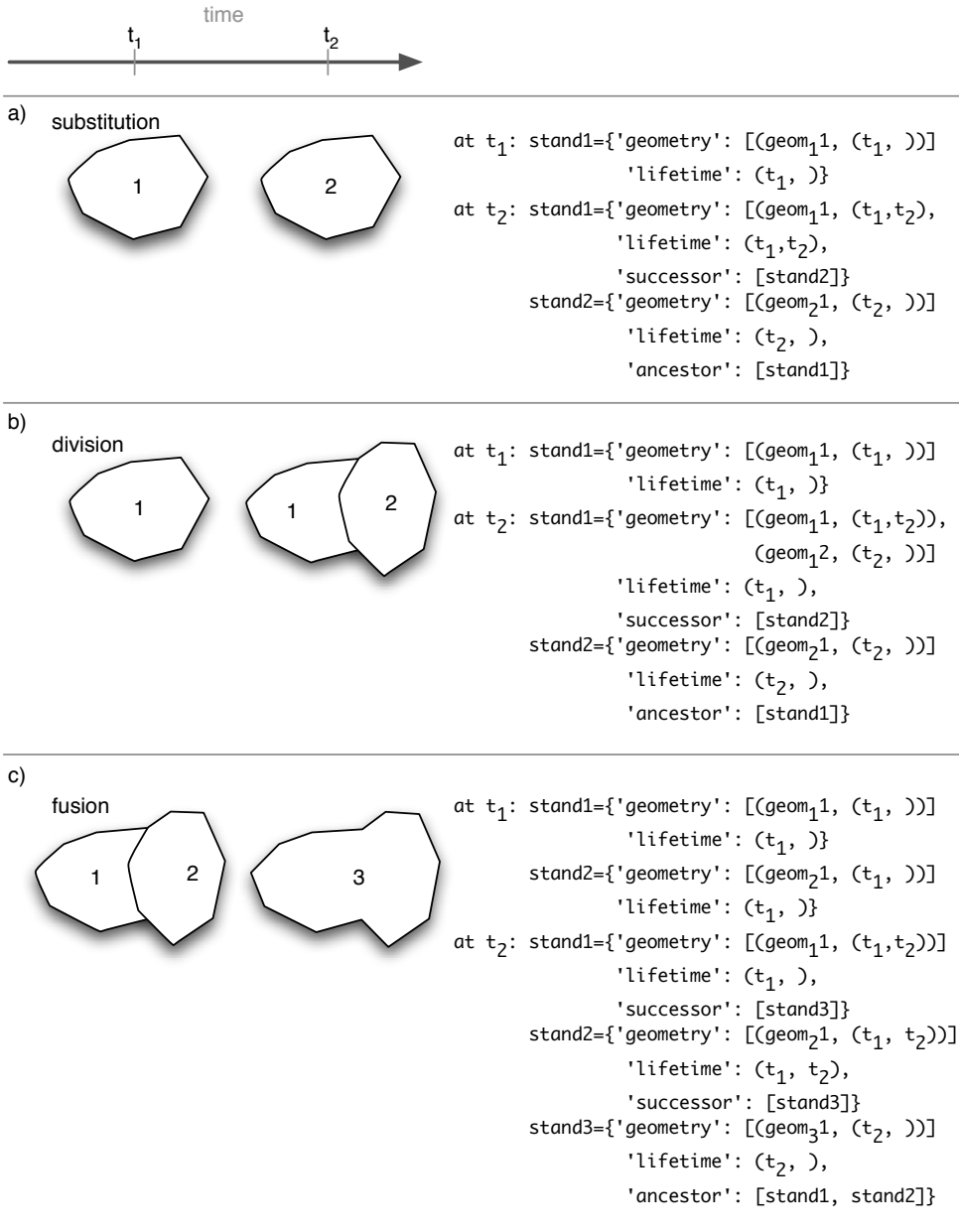
As noted earlier, the temporal object succession implementation was identical to the spatial object hierarchy implementation in the sense that object references were implicitly used to express the temporal succession both backwards and forwards in time. The sequencing operation defined in paper III could therefore be used for both spatial and temporal graph traversal.

The temporal aspect has a strong impact on the data manipulation operations. For changes in the data attributes of an object, the simple update mechanism of replacing the old attribute value with a new one must be extended to cover two new aspects. Since the data content of an object is not known at the time of creation, the object must support the addition of new attributes to it, and as a convenience function to allow for data input errors it must also, conversely, support the removal of attributes. The modification of existing attributes must support the temporal modifications triggered by a change in an attribute. In that case a period must be set for the validity of the old value and a new period initialised for that of the new value. The same applies to the current value indicator should it be utilised in implementing the data model (Fig. 10).

Another type of internal change in an object is a change in its geometry. Again three operations can be recognised, add, modify and remove, but their interpretation differs from that of attribute changes. The adding of a new geometry is equivalent to the modifying of an attribute, i.e. the period of validity of the old geometry is terminated and that of the new geometry is initialised and the latter is set as the current geometry (see stand1 in Fig. 11b). The modification of a geometry can be interpreted as a regular update operation, the old







**Figure 11.** Temporal interactions between objects and the associated changes in their spatial, lifetime and temporal succession properties. The spatial data structure is a list of tuples, each containing a data structure and a tuple of start and end dates for the geometry. The lifetime is similarly a start,end-tuple, while the temporal succession is given as lists of object identifiers.

### 4.3 Trees that once were—what do they reveal of the forest that was?

The focus in paper IV is on one of the practical hurdles in the way of a more common adaptation of the multiscale view of forests, the cost of data acquisition at or close to the individual tree level. A method was developed to create accurate reference data for a set of forest inventory grain sizes using a cost-efficient data source. The method used as its own data source semi-localised individual tree measurements from a clear cut area, as given by the harvester used in logging operations, i.e. the data are a by-product of the logging operations and as such entail negligible costs.

Individual tree locations were simulated from the location observations of the harvester at the time of cutting, using Monte Carlo simulation with two probability density functions: the distance and angle of the tree from the harvester cutting it. As a result of the simulation, each tree has a set of predicted locations, and its membership of any given subdivision of the stand can be predicted in terms of the simulation location realisations inside the subdivision as a proportion of all the simulation location realisations inside the stand.

The method was able to sharpen our view of the "forest that was" to a degree comparable with the unprocessed harvester data. The total volume estimate by species was improved by 20% and the timber assortment volume estimates by species by 17% for logs and by 35% for pulpwood. The method was progressively less able, however, to reveal accurately the properties of the forest as the observation grain size approached the single tree level. The RMSE of the volume prediction increased as a function of the decrease in subdivision area, whereas the four subdivision methods tested had no marked effect on the volume predictions. The method matches a sub-stand level, or group of trees, view of a forest, and would thus be able to produce reference data for remote sensing methods that target within-stand variation but do not go to the tree level.

## 5 Discussion

The benefit of the reference data generation method is the extreme accuracy of its individual tree volume measurements, while the drawback with regard to the single tree approach is the uncertainty of tree location, which increases the error in volume predictions the smaller the area becomes. This illustrates the asymmetric relationship between the different levels of scale: transition from a finer to a coarser level is a straightforward aggregation operation, whereas it is impossible to move in the opposite direction without prior information on the finer scale (Camossi et al. 2006). The method developed here facilitates the use, as a field source for remote sensing inventory methods or as a verification data-set, of a so far largely untapped data source involving negligible acquisition costs compared with other field data collection methods. Harvester data in itself could also serve as a source for the local adaptation of volume models. As the average throughput of a single harvester is in the range of 15–25 m<sup>3</sup>/h (Mielikäinen and Riikilä 1997), data collection is extremely efficient and could be easily automated. Accumulated over time and space, such data would permit analyses that are not possible with current data sets, such as the relationship between the form and location of trees or between log assortment and location, on both large and small geographical scales. Combined with automated data production procedures for non-timber entities (e.g. MacMillan et al. 2004) the method could assist in the production of multi-scale data-sets for forest resource management planning.

To be useful in forest management, forest resource data must be stored in such a way that it is accessible to the decision support systems used in the management planning. Data storage was successfully implemented here using all the main approaches: file, structured document and database techniques. The critical component of data storage, however, is not the actual implementation but the conceptual data model that describes the representation of the data in the information system. The search for a single, universal, representation of space-time would be a fruitless exercise, as any conceptual model for data covering both space and time is affected by the application domain (Peuquet 2002). In the domain of forest management, which embraces the timber and ecosystem aspects, the ability to utilise data on several operational scales is crucial.

The multi-scale forest resource feature hierarchy could be stored in an information system either as field-based or entity-based data. In the field-based approach a forest could be seen as a field representing the density of trees (Christakos et al. 2001), and there are efficient storage mechanisms for hierarchical, spatially nested field data, e.g. quadrees (Csillag 1997). As at the finest scale forest resource data will consist of trees, natural *bona fide* entities, the entity-based view is also justifiable, especially so as the higher levels in the data hierarchy can be seen as fiat, human-demarcated, objects. They are aggregations of *bona fide* tree entities, but their spatial demarcations are not based on *bona fide* object boundaries (Smith 2001). The choice between the two views is not a radical one, as the entity-based and field-based views are duals, i.e. a spatio-temporal phenomenon can be presented using either view (Peuquet 2002), so that continuous fields, for example, can be discretised into spatial objects whose location in space and time can be stored (Shekhar et al. 1997). The fundamental requirement, however, is that the data model should be complex enough to capture the multi-scale phenomena. Otherwise the information stored will be incomplete (Burrough 1996). The complexity of data is supported in the methods presented here by the very general definitions; essentially all that is predetermined is that the system consists of "objects of some kind" that have "properties of some kind" and that these objects can have relationships to other objects of "the same or another kind" both in space and time.

The importance of the ability to handle different kinds of objects on varying scales is emphasised when studying the interdependence of the two aspects of scale: resolution or grain, as the smallest distinguishable part of a spatial data-set, and operational scale, which refers to the scale of operation of the phenomenon under study (Lam and Quattrochi 1992). Bian (1997) suggests that the resolution of entity-based areal units overlaps with operational scale, as both are defined by biophysical features. This is beneficial, as the data can be handled directly at the correct scale for analysing the processes under scrutiny. The operational scales of the processes should be known beforehand in order to take full advantage of the resolution-operational scale match, but as this is not always the case, the model for storing and analysing the multi-scale data should allow for later aggregation to the correct operational scale if the data are recorded at a finer grain than that required for the process to be analysed. Aggregation is not unproblematic, though, as Openshaw (1984) explored in his paper on the Modifiable Areal Unit Problem. Similarly the defining of the correct resolution for the process under study is not always unproblematic, especially if there are no clearly distinguishable physical objects to be studied, but rather fiat objects originating from subjective human demarcation. Worboys (1998) touches on the same subject when analysing the effect of the schema on what is observable, maintaining that the extent of the schema specifies the size of the semantic and spatial observation window, which in turn specifies which entities are distinguishable from one another. By allowing different schemas within the confines of a very broad, generic top-level schema, the conceptual model and data proc-

essing frameworks presented here should avoid the situation where "limitations in the extent and granularity of a schema lead to incompleteness and imprecision respectively in the observation made with respect to the schema" (Worboys 1998). The schema configurability aspect of data storage and processing has an ontology aspect as well. Guarino (1997, 1998) states that as different views of reality are inherent, there can be no one unifying ontology, but hierarchies of ontologies: top-level, domain and application ontologies. The view of reality is also subject to change through time, implying that temporal data management methods should be able to accommodate ontology changes. This aspect was not explicitly studied for the data management methods presented in this work, but for data storage the implicit ontologies of different types of object, each being instances of the same base class, the temporal relationships between the objects, and the configurable semantics would suggest that changes in ontology could be accommodated. For the data processing framework an explicit but modifiable ontology for any data processing task was used, and therefore the ontology is fixed within a single processing operation, but modifications are possible between operations. This aspect has been largely ignored in recent attempts to abstract the forest resource data management framework to allow reusability of code and extensibility by means of object-oriented design (Baskent et al. 2001, Salminen et al. 2005). The commonly used approaches rely on the ability to define common base classes and their attributes and methods, but in that case the forest ontology and semantics will still be fixed inside the program code. Ontology changes at the object level pose no problems for data access, as long as the changes are implemented through the temporal relationships between objects. Changes in object attribute semantics, however, are not supported by the proposed methods. Pedersen and Jensen (1999) suggest a method for data warehousing in which analysis across changes is achieved by links between the variables that represent the "same" thing before and after the change. The semantic class could serve in the same role in the conceptual model, if it were augmented with a link property that would map between the semantically identical entities in the class. The Entity-Attribute-Value model used for object attributes in itself already naturally accommodates changes in data content by allowing changes in the attribute composition of an object without schema changes.

Peuquet (2002) argues that "fundamental issues in the development of data models that can best capture the intrinsic characteristics of geographic data thus become, first, how to overcome the difference in the nature of the phenomena being represented and that of the representation medium", where the representation medium means here a computer. She goes on to list the main differences as being the irregularity of geographical boundaries, leading to large databases, inexact and scale-dependent locational definitions, context-dependent and inexact spatial relationships and scale-specific phenomena. Even though certainly leading to even larger databases, the ability to handle multiple scales seems to be a key issue for overcoming the difference between the phenomena and their representation. The context-dependence of spatial relationships is illustrated by a study of vegetation and ecosystem classifications (Allen and Hoekstra 1990), which are hierarchical in the sense that finer classes are contained within the broader classes, but this containment does not necessarily apply to spatial nesting. To handle such situations the conceptual data model separates the explicit relationship information from the spatial information objects. This also confers the ability to model cases in which some levels in the hierarchy are operational aggregations of objects at a lower level but have the spatial resolution of a higher level in the hierarchy, e.g. the common stand-tree species stratum-tree data hierarchy, in which the stratum usually shares the geometry of the stand although it is on a lower level in the hierarchy. Even though the model employs the "forester's view of a forest" for the actual spatial

location, i.e. assuming crisp and unambiguous borders (Bennett 2001), the inexactness in spatial relationships can be implemented by extending the relationship information to include a measure of certainty regarding the existence of the link. This would allow representing spatial features using a two level hierarchy, in which at the bottom are “atomic” objects that are aggregated as top level objects. The certainty of the membership of each lower level object in each top-level object can be expressed in the certainty of the link. The certainty measure between the aggregated features and their constituting, “atomic” parts can be derived using the attribute, geometric, and neighborhood properties of the lower level features (Allan and Lowell 2002). An alternative approach to managing non-crisp or ambiguous borders would be to utilise the spatial object hierarchy to implement an egg-yolk model (Cohn and Gotts 1996) for spatial regions with indeterminate boundaries, in which a region with vague boundaries is represented as a pair of concentric regions with determinate boundaries. In that case an object could have two sub-levels, “egg” and “yolk”, with their own but interrelated geometries.

Although planning would be possible merely by constructing a representation of the current status of the forest from the multi-scale data-set, the monitoring and evaluation aspects of management would have to be bypassed, as these steps require information on the changes that have occurred. There is a clear discrepancy between the phenomena and the conceptual data models used to study changes in stand structure (Etheridge et al. 2005, Wolf 2005), forest habitat structure (Löfman and Kouki 2001, Di Orio et al. 2005, Kennedy and Spies 2005) and land cover (Wimberly and Ohmann 2004, Bender et al. 2005, Radeloff et al. 2005). These studies are based on comparisons of snapshots taken at different times, derived either from remote sensing material or inventory records. The discrepancy is unavoidable, as there are usually no other data available. The change phenomena, however, conform to a conceptual model that implies events, discrete or continuous, that drive the changes and life histories of features that record the changes at a per feature level. Worboys (2005) draws a developmental time-line for spatio-temporal models from sequences of temporal snapshots through object life histories to event chronicles, and presents an algebraic approach to events or “happenings” in order to analyse geographical phenomena over space and time. The methods presented in this work draw from all stages in such spatio-temporal models. The conceptual data model utilises life histories but extends them with a rudimentary implementation of event chronicles. The conceptual model allows for expressing event–object interactions, but lacks an event–event interaction expression mechanism. However, by treating events as analogous to objects (Galton 2001), the model immediately gains that mechanism, as a spatio-temporal object class can express object–object interaction. This could be used to shift the focus from the object-centric system state to that of process-centric analysis (Reitsma and Albrecht 2005). The handling of branching time in the data processing framework adopts the temporal snapshot sequence approach, with the notable exception that the temporal sequence does indeed record the complete life histories of objects even though not in an object-centric manner. The time concept used throughout is that of a single point, but the methods could also accommodate bi-temporal support, allowing a more fine grained approach of valid time and transactional time to the time points recorded for objects.

The objective of modern forest management, to sustain the social, ecological and economic value of forests, has implications for the extraction of timber resources. These are generally expressed as policy changes on the stand level (Lämås and Fries 1995, Karppinen 1998, Bettinger et al. 2005). The implications of these forest management policies when implemented nevertheless cover a much wider spatial range as well as a temporal one. To

evaluate the effects of proposed or implemented management policies on the social, ecological and economic processes associated with forests, one would need to be able to change the spatial and temporal scale of analysis to match that of the process under study. The lack of fully integrated support for multiple granularities of time, as opposed to the granular spatial hierarchy of objects, means, however, that a change in temporal granularity has to be implemented through granularity conversion functions (Camossi et al. 2006). This, combined with the fact that the relationships between temporal intervals (Allen 1983) are not sufficient to represent continuous change (Galton 1990), means that the conceptual model and data access methods need to be refined to fully support “why” queries in addition to “what, where and when” queries.

If the data available allow for a multi-scale approach, the analysis can either be conducted with a single model that is amended to cover all the scales relevant to the task of the analysis (Lasch et al. 2005), or by means of a modelling framework that allows for a combination of multiple scales of data and models in a coherent way. McNulty et al. (1997) tested the effect of data aggregation over three spatial scales, stand (1-10 ha), ecosystem (10-1000 ha) and region (>100,000 ha), on the predictions given by an ecosystem-level forest productivity model and found temporal and spatial aggregation to exercise a major influence on the predictions. The aggregation was performed in a non-hierarchical setting, however, where the data sources for the various model components differed greatly between the scales. Li and Reynolds (1997) promoted the approach of hierarchical modelling by studying a phenomenon on different scales in order to overcome the problem of a lack of large-scale information (with less detail) by using data and models on smaller scales (with more detail) to extrapolate to the long-term and to large spatial extent (see Moloney et al. 1992, Wu and Loucks 1995). Examples of hierarchical modelling and planning frameworks do exist covering spatial scales from sub-hectare to tens of hectares (Bettinger et al. 2005), tens of hectares to hundreds of hectares (Kurttila and Pukkala 2003) and tens of hectares to thousands of square kilometres (Seely et al. 2004). Data and implemented models are usually tightly coupled in hierarchical modelling frameworks, which is natural as they are interdependent, but at the same time the tasks that the framework is suitable for are fixed. True extensibility and adaptability would require an ability to revise the data and model components of the framework to adjust to the changing requirements for decision support. This has been acknowledged in recent developments, in which the need for an advanced data management system to support the integration of different data levels and the use of a common, extendable model base for planning has been identified (Nelson 2003, Lämås and Eriksson 2003). The diversified goals for management planning, e.g. protection of wildlife, biodiversity, scenic beauty or a reduction in water sedimentation and erosion, require not only a hierarchical approach but also an explicit implementation of spatial processing as part of the planning process. The same applies to cases in which the phenomena of interest cannot be reduced to a hierarchical representation of objects. If there is no clear containment relation between objects, they can still be modelled as independent object levels using the conceptual model. The relationship between the levels will then have to be resolved explicitly using a spatial relation operation. The direct incorporation of spatial calculations into planning models has been considered unwieldy (Baskent and Jordan 1996), but the integration of spatial properties of an object directly into the data management solution changes this. The integration of a spatial processing capability as a component of the data processing framework as presented here remains a task for future work, however. The spatial processing would also affect the combination of branching time and tracking of object life histories in the simulation results. The current time-centric model does not explicitly

support object–object interactions, as the top-level objects are processed independently of each other.

In his analysis of the representation of space and time, Galton (2001) concludes that "an effective spatio-temporal representation should be able to handle locations, times, objects and events as primitive entities, to assign attributes to any of these, and to keep track of the interdependencies amongst the various attributes assigned." Handling could be taken to mean both storage of the data, including the attributes and relationships, and access to these data. The mechanisms of access to spatio-temporal data can be divided into two approaches: programmatic and logical. Programmatic access is characterised by tight coupling between the implementation level of data storage and access to the data, while in logical access there is a layer separating the implementation level from the logical level of data storage and the access methods are matched with the logical level data storage implementation. Logical access mechanisms usually rely on a language interface, possibly augmented with a visual interface (Calcinelli and Mainguenaud 1994, Lbath et al. 1997, Elariss et al. 2006).

The main benefit of the conceptual data model presented here is its generic nature, which absorbs the changes in data content that are bound to happen over time. On the flip side of this generic nature are the complications that this causes for using the commonly prevailing logical data access mechanisms, most notably the Structured Query Language (SQL) used with relational databases. Thus, when implemented in an object-relational database, data access for the generic data model is a combination of SQL-operations and procedural program code. The circle that led from the hierarchical databases of the late 1960s to relational databases (Elmasri and Navathe 2000) is now closing with the arrival of XML, however, and we are again returning to hierarchical data access as one of the focal areas. As demonstrated, it is mainly because of its inherent extensibility that XQuery can be used as a logical access mechanism to implement data access for the generic spatio-temporal data model.

The ability to abstract data access from data storage with the help of XQuery does not come without its price. The 30 TB needed to store "all the trees in Finland" in an object database is not an unduly large figure by today's standards, although to be manageable on a personal computer, a leap of an order of magnitude is needed for the storage and processing power of a single computer. As the storage of data in an XML database consumes at least an order of magnitude more storage space, however, advances both in XML database and computer technology will be needed before the XML-XQuery combination is feasible for large data sets.

One aspect affecting the storage size needed for the spatio-temporal data-set is whether each object should be stored with its complete lineage and descendants at every successive level. As noted earlier, this would be necessary in the case of uncertain or partial links at some point in the data hierarchy. With "solid links", however, there is a choice of only storing the relationships to the object's own parents and children, in which case the data access would have to be recursive. The preferred method depends on the data access patterns, the computational cost of recursive data access and the amount of storage space saved.

All the spatial data access features studied here were based on the explicit coding of spatial relationships of objects as links between them. This was a result of the specific properties of the problem studied: time has a tendency to bring changes with it, and the different spatial scales at which a forest can be observed tend to bring with them different concepts of the forest; some corresponding strictly to real world objects and some to more human concepts that may vary between individuals. These two properties of the problem im-



ply that most of the data management solutions in wide use nowadays are of limited applicability to data management over time and over different scales.

The emphasis in data analysis is not always on tracking changes and their possible causes over time, however, or on the relationships between objects at different grain levels over time. Quite often the focus is on the various objects and their properties in a particular area at one point in time. This is where the databases in use today excel, as they are able to store and give access to data that has a constant structural content and unambiguous location. It remains a task for future work to integrate the ideas presented in this paper into the one-dimensional (e.g. Bayer and McCreight 1972) and multidimensional (e.g. Guttman 1984, Samet 1984, Hellerstein et al. 1995) index-based data access methods now in common use.

Other aspects of future work include use of the methods presented here as a basis for building a data warehouse in which spatio-temporal objects could be analysed in a by-property manner (Pedersen and Jensen 1998) instead of the by-feature approach used here. Such a data warehouse would also benefit from an interactive exploration and analysis tool (Rivest et al. 2005).

## References

- Agarwal, P. 2005. Ontological considerations in GIScience. *International Journal of Geographical Information Science* 19(5): 501–536.
- Allan, R. & Lowell, K. 2002. Managing spatial uncertainty using attribute, geometric, and neighborhood measures in an empirical rule-based model approach. *Journal of Geographical Systems* 4: 141–156.
- Allen, J.F. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM* 26(11): 832–843.
- Allen, T.F.H. & Hoekstra, T.W. 1990. The confusion between scale defined levels and conventional levels of organization in ecology. *Journal of Vegetation Science* 1(1): 5–12.
- Barbour, J., Hemstrom, M., Ager, A. & Hayes, J.L. 2005. Effects of spatial scale on the perception and assessment of risk of natural disturbance in forested ecosystems: Examples from Northeastern Oregon. *Forest Ecology and Management* 211(1-2): 210–225.
- Baskent, E.Z. & Jordan, G.A. 1996. Designing forest management to control spatial structure of landscapes. *Landscape and Urban Planning* 34(1): 55–74.
- , Wightman, R.A., Jordan, G. A & Zhai, Y. 2001. Object-oriented abstraction of contemporary forest management design. *Ecological Modelling* 143: 147–164.
- Bayer, R. & McCreight, E.M. 1972. Organization and maintenance of large ordered indexes. *Acta Informatica* 1(3): 173–189.
- Bender, O., Boehmer, H.J., Jens, D. & Schumacher, K.P. 2005. Analysis of land-use change in a sector of Upper Franconia (Bavaria, Germany) since 1850 using land register records. *Landscape Ecology* 20(2): 149–163.
- Bennett, B. 2001. What is a forest? On the vagueness of certain geographical concepts. *Topoi* 20(2): 189–201.
- Bettinger, P., Lennette, M., Johnson K.N. & Spies, T.A. 2005. A hierarchical spatial framework for forest landscape planning. *Ecological Modelling* 182(1): 25–48.

- Bian, L. 1997. Multiscale Nature of Spatial Data in Scaling Up Environmental Models. In: Quattrochi, D.A. & Goodchild, M.F. (Eds.), *Scale in Remote Sensing and GIS*, CRC Lewis Publishers, Boca Raton, FL, pp. 13–26.
- Boag, S., Chamberlin, D., Fernández, M.F., Florescu, D., Robie, J. & Siméon, J. (Eds.), 2005. XQuery 1.0: An XML Query Language. Available from: <http://www.w3.org/TR/xquery/> [Cited 18 Apr 2006].
- Bray, T., Paoli, J., Sperberg-McQueen, C.M., Maler, E. & Yergeau, F. (Eds.), 2004. Extensible Markup Language (XML) 1.0 (Third Edition). Available from: <http://www.w3.org/TR/REC-xml/> [Cited 18 Apr 2006].
- Burrough, P.A. 1996. Natural objects with indeterminate boundaries. In: *Geographic objects with indeterminate boundaries*, Burrough, P.A. & Frank, A.U. (editors), Taylor & Francis, London. pp. 3–28.
- Calcinelli, D. & Mainguenaud, M. 1994. CIGALES: A Visual Query Language for Geographical Information System: the User Interface. *International Journal of Visual Languages and Computing* 5: 113–132.
- Camossi, E., Bertolotto, M. & Bertino, E. 2006. A multigranular object-oriented framework supporting spatio-temporal granularity conversions. *International Journal of Geographical Information Science* 20(5): 511–534.
- Chamberlin, D.D. & Boyce, R.F. 1974. SEQUEL: A structured English query language. *Proceedings of the 1974 ACM SIGFIDET workshop on Data description, access and control table of contents*, pp. 249–264.
- Christakos, G., Bogaert, P. & Serre, M. 2001. *Temporal GIS: Advanced Functions for Field-Based Applications*. Springer.
- Codd, E.F. 1970. A relational model of data for large shared data banks. *Communications of the ACM* 13(6): 377–387.
- 1979. Extending the database relational model to capture more meaning. *ACM Transactions on Database Systems* 4(4): 397–434.
- Cohn, G. & Gotts, N.M. 1996. The ‘Egg-Yolk’ Representation of Regions with Indeterminate Boundaries. In: Burrough, P.A. & Frank, A.U. (Eds.), *Geographic Objects with Indeterminate Boundaries*, GISDATA series vol. 2, Taylor & Francis, pp. 171–187.
- Csillag, F. 1997. Quadtrees: Hierarchical Multiresolution Data Structures for Analysis of Digital Images. In: Quattrochi, D.A. & Goodchild, M.F. (Eds.), *Scale in Remote Sensing and GIS*, CRC Lewis Publishers, Boca Raton, FL, pp. 247–271.
- Di Orio, A.P., Callas, R. & Schaefer, R.J. 2005. Forty-eight year decline and fragmentation of aspen (*Populus tremuloides*) in the South Warner Mountains of California. *Forest Ecology and Management* 206(1-3): 307–313.
- El-Geresy, B. & Jones, C. 2000. Models and queries in a spatio-temporal GIS. In: Atkinson, P., Martin, D. (Eds.), *GIS and Geocomputation, Innovations in GIS 7*. Taylor and Francis, pp. 27–39.
- Elariss, H.E., Khaddaj, S. & Haraty, R. 2006. Towards a new visual query language for GIS. *Proceedings of the 24th IASTED international conference on Database and applications*, Innsbruck, Austria, pp. 195–202.
- Elmasri, R. & Navathe, S.B. 2000. *Fundamentals of Database Systems – 3rd ed*. Addison-Wesley.
- Etheridge, D.A., MacLean, D.A., Wagner, R.G. & Wilson, J.S. 2005. Changes in landscape composition and stand structure from 1945–2002 on an industrial forest in New Brunswick, Canada. *Canadian Journal of Forest Research* 35(8): 1965–1977.

- Ferretti, M. 1997. Forest Health Assessment and Monitoring – Issues for Consideration. *Environmental Monitoring and Assessment* 48(1): 45–72.
- Galton, A. 1990. A critical examination of Allen's theory of action and time. *Artificial Intelligence* 42(2-3): 159–188.
- 2001. Space, Time, and the Representation of Geographical Reality. *Topoi* 20(2): 173–187.
- García-Gigorro, S. & Saura, S. 2005. Forest Fragmentation Estimated from Remotely Sensed Data: Is Comparison Across Scales Possible? *Forest Science* 51(1): 51–63.
- Graf, R.F., Bollmann, K., Suter, W. & Burgmann, H. 2004. The importance of spatial scale in habitat models: capercaillie in the Swiss Alps. *Landscape Ecology* 20: 703–717.
- Griffiths, T., Fernandes, A.A. A., Paton, N., W., Barr, R. 2004. The Tripod spatio-historical data model. *Data & Knowledge Engineering* 49(1): 23–65.
- Guarino, N. 1997. Understanding, building, and using ontologies. *International Journal of Humand and Computer Studies* 46(2-3): 293–310.
- 1998. Formal ontology and information systems. *Proceedings, FOIS '98, Trento, Italy*. IOS Press, pp. 3–15.
- Güting, R.H. & Schneider, M. 1995. The International Journal on Very Large Databases 4(2): 243–286.
- Guttman, A. 1984. R-trees: a dynamic index structure for spatial searching. *Sigmod Record* 14(2): 47–57.
- Hall, J.P. 1995. Forest health monitoring in Canada: How healthy is the boreal forest? *Water, Air, & Soil Pollution* 82(1-2): 77–85.
- Hanski, I. & Ovaskainen, O. 2002. Extinction Debt at Extinction Threshold. *Conservation Biology* 16(3): 666–673.
- Hellerstein, J.M., Naughton, J.F. & Pfeffer, A. 1995. Generalized Search Trees for Database Systems. *Proceedings of the 21st International Conference on Very Large Data Bases, Zürich*, pp. 562–573.
- Horne, P., Boxall, P.C. & Adamowicz, W.L. 2005. Multiple-use management of forest recreation sites: a spatially explicit choice experiment. *Forest Ecology and Management* 207(1-2): 189–199.
- Hydrology modeling 2001. [Internet site]. Available from: <http://edndoc.esri.com/arcobjects/8.3/default.asp?url=/arcobjects/8.3/Samples/Spatial%20Analyst/Hydrology%20Modeling/HydrologyModeling.htm>. [Cited 28 Sep 2007]
- Hyvönen, P., Pekkarinen, A. & Tuominen, S. 2005. Segment-level stand inventory for forest management. *Scandinavian Journal of Forest Research* 20: 75–84.
- Iftekhar, M.S. 2005. From timber management to forest management: an initial discussion on forest management evolution. *Journal of Forest Science* 51(9): 412–429.
- Johnson, S.B. & Chatziantoniou, D. 1999. Extended SQL for Manipulating Clinical Warehouse Data. *Proceedings of the American Medical Informatics Association Symposium 1999*, pp. 819–823.
- Kalliovirta, J., Laasasenaho, J. & Kangas, A. 2005. Evaluation of the Laser-relascope. *Forest Ecology and Management* 204(2-3): 181–194.
- Kangas, A. 1997. On the prediction bias and variance in long-term growth projections. *Forest Ecology and Management* 96(3): 207–216.
- Karppinen, H. 1998. Values and Objectives of Non-industrial Private Forest Owners in Finland. *Silva Fennica* 32(1): 43–59.

- Kennedy, R.S. H. & Spies, T.A. 2005. Dynamics of hardwood patches in a conifer matrix: 54 years of change in a forested landscape in Coastal Oregon, USA. *Biological Conservation* 122(3): 363–374.
- Köchli, D.A. & Brang, P. 2005. Simulating effects of forest management on selected public forest goods and services: A case study. *Forest Ecology and Management* 209(1-2): 57–68.
- Korpela, I. & Tokola T. 2006. Potential of aerial image-based monoscopic and multiview single-tree forest inventory – a simulation approach. *Forest Science* 52(2):136–147.
- Kurki, S., Nikula, A., Helle, P. & Lindén, H. 2000. Landscape Fragmentation and Forest Composition Effects on Grouse Breeding Success in Boreal Forests. *Ecology* 81(7): 1985–1997.
- Kurttila, M. & Pukkala, T. 2003. Combining holding-level economic goals with spatial landscape-level goals in the planning of multiple ownership forestry. *Landscape Ecology* 18: 529–541.
- Lam, N.S.-N. & Quattrochi, D.A. 1992. On the issues of scale, resolution, and fractal analysis in the mapping sciences. *Professional Geographer* 44(88): 88–98.
- Lämås, T. & Fries, C. 1995. Emergence of a biodiversity concept in Swedish forest policy. *Water, Air, & Soil Pollution* 82(1-2): 57–66.
- & Eriksson, L.O. 2003. Analysis and planning systems for multiresource, sustainable forestry: the Heureka research programme at SLU. *Canadian Journal of Forest Research* 33: 500–508.
- Langran, G. 1991. Time in Geographic Information Systems. *Technical Issues in GIS Series*. Taylor & Francis.
- Lasch, P., Badeck, F.-W., Suckow, F., Lindner, M. & Mohr, P. 2005. Model-based analysis of management alternatives at stand and regional level in Brandenburg (Germany). *Forest Ecology and Management* 207(1-2): 59–74.
- Lbath, A., Aufrane-Portier, M.A. & Laurini, R. 1997. Using a visual language for the design and query in GIS customization. *Proceedings of the 2nd International Conference on Visual Information Systems*, San Diego, USA, pp. 197–204.
- Levin, S.A. 1992. The problem of pattern and scale in ecology. *Ecology* 73(6): 1943–1967.
- Li, H. & Reynolds J.F. 1997. Modelling Effects of Spatial Pattern, Drought, and Grazing on Rates of Rangeland Degradation: A Combined Markov and Cellular Automaton Approach. In: Quattrochi, D.A. & Goodchild, M.F. (Eds.), *Scale in Remote Sensing and GIS*, CRC Lewis Publishers, Boca Raton, FL, pp. 211–230.
- Löfman, S. 2006. Changes in forest landscape structure in southern Finland in the late 1900's. *Dissertationes Forestalis* 32.
- & Kouki, J. 2001. Fifty years of landscape transformation in managed forests in Southern Finland. *Scandinavian Journal of Forest Research* 16: 44–53.
- MacMillan, R.A., Jones, R.K. & McNabb, D.H. 2004. Defining a hierarchy of spatial entities for environmental analysis and modeling using digital elevation models (DEMs). *Computers, Environment and Urban Systems* 28: 175–200.
- Maltamo, M., Mustonen, K., Hyypä, J., Pitkänen, J. & Yu, X. 2004. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. *Canadian Journal of Forest Research* 34(9): 1791–1801.
- Maser, C., Bormann, B.T., Brookes, A.R., Kiester, A.R. & Weigand, J.F. 1994. Sustainable forestry through adaptive ecosystem management is an open-ended experiment. In: Maser, C. (Ed.), *Sustainable forestry: Philosophy, science, and economics*, St. Lucie Press, Delray Beach, FL, pp. 304–340.

- McNulty, S.G., Vose, J.M. & Swank, W.T. 1997. Scaling Predicted Pine Forest Hydrology and Productivity Across Southern United States. In: Quattrochi, D.A. & Goodchild, M.F. (Eds.), *Scale in Remote Sensing and GIS*, CRC Lewis Publishers, Boca Raton, FL, pp. 187–209.
- Mendoza, M.A., Fajardo, J.J. & Zepeta, J. 2005. Landscape based forest management, a real world case study from Mexico. *Forest Ecology and Management* 209(1-2): 19–26.
- Mielikäinen, K. & Riikilä, M. 1997. Kannattava puuntuotanto. *Metsälehti kustannus*. (In Finnish)
- Moloney, K.A., Levin, S.A., Chiariello, N.R. & Buttel, L. 1992. Pattern and scale in a serpentine grassland. *Theoretical Population Biology* 41(3): 257–276.
- Montes, F., Sánchez, M., del Río, M. & Cañellas, I. 2005. Using historic management records to characterize the effects of management on the structural diversity of forests. *Forest Ecology and Management* 207(1-2): 279–293.
- Mountrakis, G., Agouris, P. & Stefanidis, A. 2002. A differential spatio-temporal model: primitives and operators. In: *Advances in Spatial Data Handling*, 10th International Symposium on Spatial Data Handling, July 9–12, Ottawa, Canada, Heidelberg: Springer-Verlag, pp. 255–268.
- Moore, C.T., Conroy, M.J., Boston, K. 2000. Forest management decisions for wildlife objectives: system resolution and optimality. *Computers and Electronics in Agriculture* 27(1-3): 25–39.
- Nams, V.O., Mowat, G. & Panian, M.A. 2006. Determining the spatial scale for conservation purposes – an example with grizzly bears. *Biological Conservation* 128(1): 109–119.
- Nelson, J. 2003. *Forest Planning Studio (FPS)-ATLAS Program: Reference Manual Version 6*. Faculty of Forestry, University of BC, Vancouver, BC, 104 pp. Available from: [http://www.forestry.ubc.ca/atlas-simfor/extension/docs.html#FPS\\_2003](http://www.forestry.ubc.ca/atlas-simfor/extension/docs.html#FPS_2003). [Cited 28 Sep 2007]
- Openshaw, S. 1984. The Modifiable Areal Unit Problem. *CATMOG*, 38, Geobooks, Norwich, England.
- Oracle Berkeley DB 2007. [Internet site]. Available from: <http://www.oracle.com/database/berkeley-db/index.html>. [Cited 28 Sep 2007]
- Pedersen, T.B. & Jensen C.S. 1998. Research issues in clinical data warehousing. *Proceedings of the 10th International Conference on Scientific and Statistical Database Management*, Capri, Italy, pp. 43–52.
- & Jensen, C.S. 1999. Multidimensional Data Modeling for Complex Data. *15th International Conference on Data Engineering (ICDE'99)*, pp. 336–345.
- Pekkarinen, A. 2002. A method for the segmentation of very high spatial resolution images of forested landscapes. *International Journal of Remote Sensing* 23(14): 2817–2836.
- Pennanen, J., Greene, D.F., Fortin, M.-J. & Messier, C. 2004. Spatially explicit simulation of long-term boreal forest landscape dynamics: incorporating quantitative stand attributes. *Ecological Modelling* 180(1): 195–209.
- Peuquet, D.J. 2002. *Representations of Space and Time*. The Guilford Press, New York London.
- & Duan, N. 1995. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. *International Journal of Geographical Information Systems* 9 (1), 7–24.
- Pitkänen, R. 2006. Monilähteinen metsien inventointi. *Metsäntutkimus* 4/2006: 10–16. (In Finnish)

- Radeloff, V.C., Hammer, R.B. & Stewart, S.I. 2005. Rural and Suburban Sprawl in the U.S. Midwest from 1940 to 2000 and Its Relation to Forest Fragmentation. *Conservation Biology* 19(3): 793–805.
- Rauscher, H.M. & Reynolds, K. 2005. Decision-support systems for forest management. *Computers and Electronics in Agriculture* 49: 1–5.
- Reitsma, F. & Albrecht, J. 2005. Implementing a new data model for simulating processes. *International Journal of Geographical Information Science* 19(10): 1073–1090.
- Reynolds, K.M. 2005. Integrated decision support for sustainable forest management in the United States: Fact or fiction? *Computers and Electronics in Agriculture* 49(1): 6–23.
- & Hessburg, P.F. 2005. Decision support for integrated landscape evaluation and restoration planning. *Forest Ecology and Management* 207(1-2): 263–278.
- Rivest, S., Bédard, Y., Proulx, M.-J., Nadeau, M., Hubert, F. & Pastor, J. 2005. SOLAP technology: Merging business intelligence with geospatial technology for interactive spatio-temporal exploration and analysis of data. *ISPRS Journal of Photogrammetry & Remote Sensing* 60(1): 17–33.
- Salminen, H., Lehtonen, M. & Hynynen, J. 2005. Reusing legacy FORTRAN in the MOTTI growth and yield simulator. *Computers and Electronics in Agriculture* 49: 103–113.
- Samet, H. 1984. The Quadtree and Related Hierarchical Data Structures. *ACM Computing Surveys* 16(2): 187–260.
- Saura, S. & Martínez-Millán, J. 2000. Landscape patterns simulation with a modified random clusters method. *Landscape Ecology* 15(7): 661–678.
- Seely, B., Nelson, J., Wells, R., Peter, B., Meitner, M., Anderson, A., Harshaw, H., Shepard, S., Bunnell, F.L., Kimmins, H. & Harrison, D. 2004. The application of a hierarchical, decision-support system to evaluate multi-objective forest management strategies: a case study in northeastern British Columbia, Canada. *Forest Ecology and Management* 199(2-3): 283–305.
- Shekhar, S., Coyle, M., Goyal, B., Liu, D.-R. & Sarkar, S. 1997. Data models in geographic information systems. *Association for Computing Machinery. Communications of the ACM* 40(4): 103–111.
- Smith, B. 2001. Fiat Objects. *Topoi* 20: 131–148.
- Stendahl, J. & Dahlin, B. 2002. Possibilities for Harvester-based Forest Inventory in Thinnings. *Scandinavian Journal of Forest Research* 17: 548–555.
- Thierry, B. & Lowell, K. 2001. An Uncertainty-Based Method of Photointerpretation. *Photogrammetric Engineering and Remote Sensing* 67(1): 65–72.
- Turner, M.G., O'Neill, R.V., Gardner, R.H. & Milne, B.T. 1989. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* 3: 153–162.
- Wachowicz, M. 1999. *Object-oriented Design for Temporal GIS*. Taylor & Francis.
- Walters, C.J. 1986. *Adaptive Management of Renewable Resources*. MacMillan Publishing Company, New York.
- Wiens, J.A. 1989. Spatial Scaling in Ecology. *Functional Ecology* 3(4): 385–397.
- Wimberly, M.C. & Ohmann, J.L. 2004. A multi-scale assessment of human and environmental constraints on forest land cover change on the Oregon (USA) coast range. *Landscape Ecology* 19(6): 631–646.
- Wolf, A. 2005. Fifty year record of change in tree spatial patterns within a mixed deciduous forest. *Forest Ecology and Management* 215: 212–223.

- Worboys, M.F. 1994. Unifying the spatial and temporal components of geographical information. In: *Advances in GIS Research, Sixth International Symposium on Spatial Data Handling (SDH'94)*, pp. 505–517.
- 1998. Computation with imprecise geospatial data. *Computers, Environment and Urban Systems* 22(2): 85–106.
- 2005. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science* 19(1): 1–28.
- Wu, J. & Loucks, O.L. 1995. From Balance of Nature to Hierarchical Patch Dynamics: A Paradigm Shift in Ecology. *The Quarterly Review of Biology* 70(4): 439–466.
- Yuan, M. 1999. Use of a three-domain representation to enhance GIS support for complex spatiotemporal queries. *Transactions in GIS* 3(2), 137–159.
- Yue, T.X., Liu, J.Y., Li, X.Q., Chena, S.Q., Maa, S.N., Tian, Y.Z. & Gec, F. 2005. Considerable effects of diversity indices and spatial scales on conclusions relating to ecological diversity. *Ecological Modelling* 188: 418–431.
- Zenner, E.K. 2005. Investigating scale-dependent stand heterogeneity with structure-area-curves. *Forest Ecology and Management* 209(1-2): 87–100.
- zlib 1996. [Internet site]. Available from: <http://www.zlib.net/>. [Cited 28 Sep 2007].